DEVELOPMENT OF THE SECONDARY-BIOLOGY CONCEPT INVENTORY (S-BCI): A STUDY OF CONTENT AND CONSTRUCT VALIDATION

Andria STAMMEN
The Ohio State University, U.S.A.

Deb LAN
The Ohio State University, U.S.A.

Anita SCHUCHARDT
University of Pittsburgh, U.S.A.

Kathy MALONE
The Ohio State University, U.S.A.

Lin DING
The Ohio State University, U.S.A.

Zakee SABREE
The Ohio State University, U.S.A.

William BOONE
Miami University, U.S.A.

ABSTRACT: This project aims to develop a measurement tool for assessing the conceptual understanding of secondary grade-level biology students (ages 11 to 18) that is reliable and valid. The study reported here describes the validity assessment of Secondary Biology Concept Inventory (S-BCI). A pool of assessment tasks was designed to target major biology constructs. The assessment items' answer stems were developed to include distractors representing students' alternative conceptions obtained from literature and student interviews. The validation stage of the S-BCI development involved an iterative revision and review process to help establish sufficient S-BCI content and construct validity. This stage included (i) student interviews and (ii) multi-expert panel critique. Based on the results of the aforementioned analyses, assessment items were proven to be valid where included on the S-BCI.

Key words: biology, secondary education, concept inventory, and alternative conceptions

INTRODUCTION

Concept inventories (CIs) are research-based measurement instruments used for assessing student understanding of concepts (Hestenes et al., 1992). These standardized selected response tests can be useful tools in measuring what students have learned in secondary science. Several existing CIs target tertiary-level conceptual understanding of specific topics in biology such as natural selection, cell division, genetics, and osmosis and diffusion (Anderson et al., 2002; Elrod, 2008; Nehm and Reilly, 2007; Odom and Barrow, 1995; Parker et al., 2008; Williams et al., 2008; and Wilson et al., 2006). Additionally, the college-level Biology Concept Inventory (BCI) includes the major concepts covered in a first-year undergraduate biology course. However, the BCI's validation process included college-level students and not secondary-level students (Klymkowsky & Garvin-Doxas, 2008). Although there are several existing CIs related to biology concepts, there is no fully developed CI available that collectively measures the major concepts covered in secondary biology classrooms. Thus, this study aims to develop a measurement tool for assessing the conceptual understanding of secondary grade-level biology students (grades 7 to 12) that is reliable and valid. In this paper, we describe the Secondary-Biology Concept Inventory (S-BCI) and its development and validation.

METHODS

Our goal in developing the S-BCI was to design a concept inventory grounded in student understanding that would be able to measure the thinking of a large, diverse sample of secondary-level biology students. The instrument needed to produce both reliable and valid data while distinguishing among students with different levels of secondary-level biology knowledge. With these goals in mind, the S-BCI was developed based on alternative conceptions identified from both a literature review and student interviews (*N*=15). Pools of assessment items were created for consideration (*N*=61) and then modified based on feedback from Expert Panels, composed of biology content experts as well as master level biology high school teachers.

S-BCI Constructs and Item Development

The S-BCI was designed to assess secondary-level students' understanding of core concepts (Table 1). These core concepts were identified by surveying a panel of biology teachers and experts about which concepts represent the fundamental models in the field of biology and are taught at the secondary-level. The following five core concepts emerged from this survey: (i) evolution and diversity, (ii) population interactions, (iii) growth and reproduction, (iv) inheritance, and (v) energy and matter. Each core concept aims to address an essential question (Table 1).

Table 1. Essential questions associated with the core concepts

Core concepts in S-BCI	Essential questions
CC1. Evolution and diversity	How and why do populations change over time?
CC2. Population interactions	How and why do populations in a system interact with other populations over time?
CC3. Growth and reproduction	How is information preserved during reproduction while still produce the variation observed in life?
CC4. Inheritance	**How are traits passed from parents to offspring?**
CC5. Energy and matter	How and why do energy and matter transfer within and across systems?

An average of 12 single-tiered items were written or adapted from other assessments targeting the core concepts associated with each model (Table 2). A total of 61 selected response items were developed. Each item was comprised of question stem and four to seven possible responses. Many of the distractor responses represented alternative conceptions identified by practitioner observations and empirical research.

Table 2. Core concepts in S-BCI

Core concepts in S-BCI	Total number of questions
CC1. Evolution and diversity	13
CC2. Population interactions	12
CC3. Growth and reproduction	13
CC4. Inheritance	11
CC5. Energy and matter	13
TOTAL	61

Validation

The validation stage of the S-BCI development involved an iterative revision and review process to help establish sufficient S-BCI content and construct validity. This stage included (i) student interviews, (ii) student questionnaires, and (iii) multi-expert panel critique. Based on the responses from the multi-expert panel review, student questionnaires, and student interviews, the S-BCI items were revised.

The multi-expert panel critique stage entailed receiving feedback from two distinct panels: (i) Biology Expert Panel and (ii) High School Expert Panel. The Biology Expert Panel was comprised of five staff and faculty members representing three distinct universities. The High School Expert Panel included eight teachers representing eight public and private high schools. These eight teachers had on average 17 years of experience. These panels critically analyzed the S-BCI for factual/conceptual accuracy, diagrammatic accuracy, alternative conception alignment, and the age-appropriateness of item structure and content including readability metrics.

The student interview stage involved students who were enrolled in undergraduate courses at a public large university (*N*=7) and secondary students enrolled in a biology course at a public high school (*N*=8), respectively. These public learning institutions are both located in the Midwestern United States. Using a "think aloud" interview structure, students, both undergraduate and high school, were asked to explain their understanding of each item's question stem and answer stems.

Additionally, secondary students completed an open-ended questionnaire which asked students to explain their understanding of the items' question stems and common vocabulary terms used in the S-BCI (*N*=73). These interview and questionnaire data were explored for alternative conceptions lacking an empirical presence in literature, in addition to analyzing if students chose correct answers for the correct reasons. Based on the responses from the multi-expert panel review and student interviews and questionnaires, the S-BCI items were revised.

RESULTS

In validating the S-BCI, the first stage of developing content and construct validity included student interviews. From these student interviews, items that were designated as having 'validity concerns' were edited for student questionnaires. Generally, the validity concerns of the assessment items fell under 3 categories: (1) confusion about the wording and problems with complex terms, (2) a lack of understanding of figures associated with the question, and (3) unanticipated alternative conceptions. Each category is briefly defined and then an exemplar is provided to illustrate the S-BCI validation process.

The first category, confusion about the wording and problems with complex terms, is associated with vocabulary within the question stem that students did not understand. These words were first identified in the undergraduate student interviews by students either asking for clarification of the term or a lack of understanding of the terms' definition after further questioning.

The second category, a lack of understanding of figures associated with the question, was discovered during undergraduate interviews and further explored in the high school interviews. During the interviews, each figure was evaluated on two criteria: (i) whether the figure was appropriate for the question; and (ii) whether the figure was necessary to answer the question. Figures that were highly complex and/or not descriptive enough were revised following the undergraduate interviews. The revised figures were then shown during interviews to high school students and further revised when necessary.

The third category, unanticipated alternative conceptions arose from adjustments made to the question stem or response options as a result of alternative conceptions students had that were discovered through the interviews. Exemplar for each category and its progression from expert panel through interviews is described below. These exemplars represent examples of how items were modified during the S-BCI validation process.

Exemplar I: Confusion about Wording

The first exemplar assessment item represents an example of a question that was identified as having validity concerns during the expert panel review and student interviews because of lexicon complexity. The original item (Figure 1) was developed for the Dominance Concept Inventory (Abraham, Perez & Price, 2014). This question was incorporated into the S-BCI because the item aligned with the S-BCI's Inheritance Core Concept (Tables 1 and 2). Furthermore, this task targets common alternative concepts held by some secondary-level students. For example, if a student selects distractor B, then the students may have the alternative conception that within a population, the selective advantage of a particular phenotype is determined by the phenotype's impact on survival and reproduction (Abraham, Perez & Price, 2014).

Figure 1. Original question (Abraham, Perez & Price, 2014)

A rose population has two alleles of a gene for thorn length. Long thorns help protect the roses from herbivory by deer. Allele H1 codes for long thorns, while allele H2 codes for short thorns. Given this information, please indicate which of the following a biologist would infer about the mode of inheritance for allele H2?

a) It is dominant.
b) It is recessive.
c) It is co-dominant.
d) It is impossible to determine.

The original question (Figure 1) was reviewed by both the Biology Expert Panel and High School Expert Panel. The expert panel review data suggested that description of the allele variants (i.e. allele H1 codes and allele H2 codes) may cause student confusion. The term 'herbivory' was also identified as a term that may lead to student comprehension issues. Therefore, this question was edited to reduce student confusion towards science specific lexicon (Figure 2).

Figure 2. Edited question after expert panel review

In a rose population, there are two variants for thorn length, short thorns and long thorns. Long thorns help protect the roses from being eaten by deer. Given this information, please indicate which of the following a biologist would infer about the mode of inheritance for short thorns?

a) It is dominant.
b) It is recessive.
c) It is co-dominant.
d) It is impossible to determine.

The assessment item in Figure 2 was further evaluated for validity during the undergraduate interview stage. During these interviews, the students indicated that the phrase ‘mode of inheritance’ could be reworded in order to reduce terminology confusion. As a result, the phase ‘mode of inheritance’ in the original question (Figures 1 and 2) was replaced with the phrase ‘the way [thorns] are inherited’ (Figure 3). Furthermore, the answer stems were expanded to include descriptions of student reasoning related to her/his conceptual understanding. That is, if a student selects choice A, then that student is likely to have the alternative conception that dominant traits provide an adaptive advantage.

Figure 3. Edited question after undergraduate interviews

In a rose population, there are two variants for thorn length, short thorns and long thorns. Long thorns help protect the roses from being eaten by deer. Given this information, please indicate which of the following a biologist would infer about the way short thorns are inherited?

a) It is a dominant inheritance pattern because short thorns have an adaptive advantage.
b) It is a recessive inheritance pattern because short thorns are more widespread in the population.
c) It is a co-dominant inheritance pattern because both long and short thorns are found in the population.
d) It is impossible to determine.

Exemplar II: Figures Associated with the Question

The second exemplar assessment item represents a question that was identified as having validity concerns related to figure representation during the expert panel review and student interviews. The original item (Figure 4) was developed by the research team for S-BCI because the item aligned with the S-BCI’s Evolution and Diversity Core Concept (Tables 1 and 2). The assessment item targets common alternative concepts held by some secondary-level students. For example, if a student selects distractor D, then the students may have the alternative conception that natural selection is only related to the survival of the strongest organisms in a population.

Figure 4. Original question

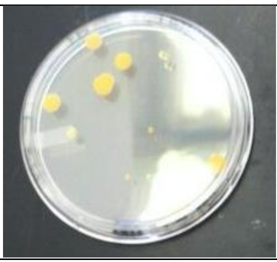
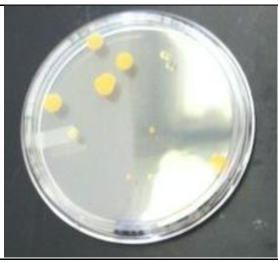
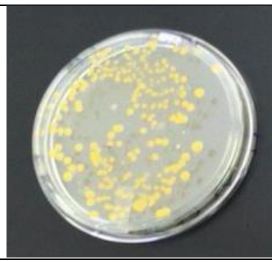
A biologist has been growing a population of bacteria on a growth media containing an antibiotic for 2 days and then switching the bacteria to media without the antibiotic for 2 days. The biologist noticed that initially very few bacteria survived (i.e., were resistant to the antibiotic), but now almost 100% of the bacteria survive. He proposes that

a) The environment (the growth media with the antibiotic) caused the bacteria to become resistant to the antibiotic. Each generation more bacteria changed and so now more survive.
b) The environment (the growth media with the antibiotic) caused the bacteria to become resistant to the antibiotic. These bacteria survived and increased their reproduction more than nonresistant bacteria and so now more survive.
c) Initially, some bacteria were resistant to the antibiotic and some weren’t. The environment (the growth media with the antibiotic) allowed those that were resistant to survive and reproduce better than the nonresistant bacteria and so now more survive.
d) Initially, some bacteria were stronger than the other bacteria. The environment (the growth media with the antibiotic) allowed those that were stronger to survive and reproduce better than the weaker ones and so now more survive.

During the next phase of the validation process for this item, the expert panels’ suggestion that the context of the question stem assumed knowledge of laboratory procedures and equipment/materials in addition to understanding natural selection in bacteria populations was incorporated into the modification of this question. Specifically, the panel proposed that including a pictorial representation would help students conceptualize the laboratory procedures and equipment/materials associated with this item (Figure 5). For the undergraduate interviews, the question was edited to include more information within the table including the images. During the undergraduate interviews, there was confusion surrounding the abundance of information that was initially given in paragraph form. Several of the undergraduate students found the information redundant. All of the information about the biologists' procedures were summarized in the table, rather than in a paragraph within the question stem. The term "media/medium" was removed and replaced with "dish" because the word "media/medium" directed students towards thinking about digital media rather than growth media. After the undergraduate interviews the question was edited as shown in Figure 5.

Figure 5. Edited question after undergraduate interviews

A biologist is conducting an experiment using bacteria. See the figure below for her procedure:

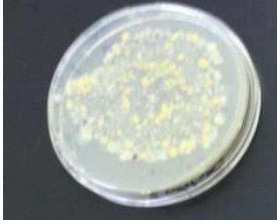
Day 1	**Day 2**	**Day 3**	**Day 4**
			
Bacteria (small dots) in a dish containing nutrients and an antibiotic.	Bacteria in a dish containing nutrients and an antibiotic. Some of this bacterium is moved to the new dish on day 3.	Moved bacteria in a dish containing no antibiotic.	Bacteria in a dish containing no antibiotic.

After a year of doing this rotation with this population of bacteria, the biologist noticed that, while initially very few bacteria survived, now almost 100% of the bacteria survive when placed in a dish with nutrients and an antibiotic. What might she conclude:

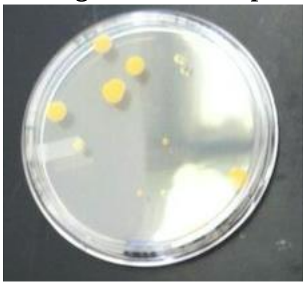
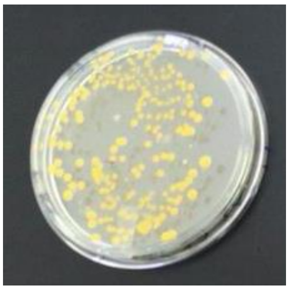
a) The environment (with the antibiotic) caused the bacteria to become resistant to the antibiotic. Each generation more bacteria changed and so now more survive.
b) The environment (with the antibiotic) caused the bacteria to become resistant to the antibiotic. These bacteria survived and increased their reproduction more than nonresistant bacteria and so now more survive.
c) Initially, some bacteria were resistant to the antibiotic and some weren't. The environment (with the antibiotic) allowed those that were resistant to survive and reproduce better than with the nonresistant bacteria and so now more survive.
d) Initially, some bacteria were stronger than the other bacteria. The environment (with the antibiotic) allowed those that were stronger to survive and reproduce better than the weaker ones and so now more survive.

During the high school interviews, it was found that high school students continued to find the images as confusing. Students associated the same growth plate across all four days without understanding the different types of combinations of nutrient and antibiotic in each plate. Therefore, after high school interviews the figure was redesigned so that the days of the dish containing both nutrients and antibiotic from those of the dish containing nutrient but no antibiotic were separated from each other. The edited image for the question can be found below in Figure 6. The question stem and choices for response remained the same.

Figure 6. Edited question after high school interviews

<table>
<tr><td colspan="2">EXPERIMENT 1: Bacteria is grown in the initial growth plate. The biologist adds antibiotic to the plate and checks the plate after 1 week. The bacterial growth is shown in the second plate below.</td></tr>
<tr><td>Initial Growth

View of plate from initial growth</td><td>Growth after 1 week

View of plate 1 week after adding antibiotic</td></tr>
</table>

<table>
<tr><td colspan="2">EXPERIMENT 2: The biologist removes bacteria from the second plate (above) and puts it on a new plate (without antibiotics) the initial growth and growth after 1 week are shown below.</td></tr>
<tr><td>Initial growth on new plate

View of bacteria on new plate (without antibiotics)</td><td>Growth after 1 week

View of bacteria after 1 week on new plate (without antibiotics)</td></tr>
</table>

Exemplar III: Unanticipated Alternative Conceptions

The final exemplar assessment item represents an example of a question displaying unanticipated alternative conceptions held by students. This validity concern emerged only during student interviews, and thus was not identified during the expert panel critique. The original item (Figure 7) was adapted from the Diagnostic Question Clusters related to Tracing Matter in Dynamic Systems assessment (Wilson, Anderson, Heidemann, Merrill, Merritt, Richmond, Silbey, & Parker, 2006). This question was incorporated into the S-BCI because the item aligned with the S-BCI's Energy and Matter Core Concept (Tables 1 and 2).

Figure 7. Original question (Wilson et al., 2006)

A mature maple tree can have a mass of more than a ton (dry mass, after removing the water), yet it starts from a seed that weighs less than 2 grams. Which of the following processes contributes the most to this huge increase in biomass?

a. Absorption of mineral substances from the soil via the roots.
b. Absorption of organic substances from the soil via the roots.
c. Absorption of carbon dioxide into molecules by green leaves.
d. Absorption of water from the soil into molecules by green leaves.
e. Absorption of solar radiation from the sun by green leaves.

While no edits were made following undergraduate student interviews, high school student interviews indicated that the question stem needed to be revised. Removal of the phrase 'maple tree' from the original question stem occurred because it was discovered during the high school student interviews that a specific type of plant guided the student into thinking about more complex structures that contribute to the increase in weight. Using 'maple tree' rather than general 'plant seed' directed students towards thinking of possible mechanisms for increase in weight that are maple tree specific. After the high school interviews, the question was edited as shown in Figure 8.

Figure 8. Edited question after high school interviews

A scientist weighed a plant seed and found that it was less than 1 gram. She planted the seed. When the seed was a height of 10 meters she weighed, it using a really big crane. She found it weighed over a ton. What do you think contributes most to this huge increase in weight?

a. Absorption of mineral substances from the soil via the roots.
b. Absorption of organic substances from the soil via the roots.
c. Absorption of carbon dioxide into molecules by leaves.
d. Absorption of water from the soil into molecules by leaves.
e. Absorption of solar radiation from the sun by leaves.

CONCLUSION

After the expert panels and student interviews, both an undergraduate and high school, the refined S-BCI questions totaled 52. Nine of the original 61 questions were considered invalid. The remaining 52 S-BCI questions will be moved forward in large scale quantitative testing. Within the interview stage both undergraduate and high school students understood the questions with a mix of students conceptually targeting, both alternative and accepted conceptions, indicating that the items on the S-BCI vary in difficulty and included a range of conceptions within student understanding. Table 3 (below) shows the finalized break down of core concepts embedded within the S-BCI that will be pilot tested in a wide-scale quantitative study.

Table 3. Core concepts in s-bci post panel review and interviews

Core concepts in S-BCI	Total number of questions
CC1. Evolution and diversity	13
CC2. Population interactions	9
CC3. Growth and reproduction	11
CC4. Inheritance	6
CC5. Energy and matter	13
TOTAL	52

Based on the results from the expert panel, undergraduate interviews, and high school interviews, we identified the items with validity concerns and were able to edit them in an iterative cycle in order to ensure that, at the end of this stage, the S-BCI would be ready to be pilot tested in a wide-scale quantitative study. In the wide scale, quantitative study, we will work to determine if the S-BCI is reliable.

FUTURE STEPS

The next step in the development of the S-BCI is the reliability testing. This will include a sample of students totaling over 1800 students in grades 8 through 12. The S-BCI items will be administered to students enrolled in science courses at seven public high schools in five states. The students are from rural, suburban and urban areas of the United States. Each student received a test with 34 questions. The number of items given to each student ensured that each item would be given to multiple students in order to obtain discrimination data while also allowing for questions that would be taken by all students.

Quantitative analysis on the S-BCI will analyze item difficulty levels, discrimination indices, point bi-serial coefficients, and Ferguson's delta. These are separated into individual item analysis and whole test analysis. Individual test analysis includes item difficulty levels, discrimination indices, and point bi-serial coefficients. Ferguson's delta is a whole test reliability analysis.

REFERENCES

Abraham, J. K., Perez, K. E., & Price, R. M. (2014). The Dominance Concept Inventory: A Tool for Assessing Undergraduate Student Alternative Conceptions about Dominance in Mendelian and Population Genetics. *CBE Life Sciences Education*, *13*(2), 349–358. http://doi.org/10.1187/cbe.13-08-0160

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the Conceptual Inventory of Natural Selection. *Journal of Research in Science Teaching*, 39: 952–978.

Elrod, S. (2008). Genetics concept inventory (GenCI) development. 2008. Paper presented at Conceptual Assessment in Biology Conference II; 3–8 January, Asilomar, California. (25 September 2008; *http://bioliteracy.net/CABS.html*).

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *Physics Teacher*, 30: 141–158.

Klymkowsky, M. W., & Garvin-Doxas, K. (2008). Recognizing students' misconceptions through Ed's Tools and the Biology Concept Inventory. PloS Biology 6: e3. doi: 10.1371/journal.pbio.0060003

Nehm. R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57: 263–272.

Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32: 45–61.

Parker J., et al. (2008). Frameworks for reasoning and assessment in Mendelian genetics. Paper presented at Conceptual Assessment in Biology Conference II; 3–8 January, Asilomar, California. (25 September 2008; *http://bioliteracy.net/CABS.html*).

Williams, K., Fisher, K., & Anderson, D. (2008). Using diagnostic test items to assess conceptual understanding of basic biology ideas: A plan for programmatic assessment. Paper presented at the Conceptual Assessment in Biology Conference II; 3–8 January, Asilomar, California. (25 September 2008; *http://bioliteracy.net/CABS.html*).

Wilson, C. D., Anderson, C.W., Heidemann, M., Merrill, J. E., Merritt, B. W., Richmond, G.,

Silbey, D. F., & Parker, J. M. (2006). Assessing students' ability to trace matter in dynamic systems in cell biology. *Life Sciences Education*, *5*: 323–331.